
CHRISTOPHE BENAVENT

COMMENT L'APPRENTISSAGE
ARTIFICIEL CHANGE
NOTRE MONDE

39

*B*ig Data, IA, plateformes, transformation digitale, blockchain, métavers, le lexique des technologies de l'information foisonne, excitant l'imagination, promettant des miracles, se nourrissant de fantasmes et d'inquiétudes, et au premier plan la peur que la machine prenne le pouvoir. Sur un plan pragmatique, le débat se pose en termes de dualité, entre substitution et complémentarité. Les optimistes voient dans ces techniques le moyen d'augmenter les capacités humaines, les pessimistes les considèrent comme une source d'aliénation.

Plutôt que de partir du jugement moral et de sa balance entre le bien et le mal, pour mieux comprendre comment l'apprentissage artificiel, qui est une expression plus juste de la nature de ces technologies, affecte nos sociétés, appuyons-nous sur une idée simple : notre vie sociale et économique, nos modes de coordination et d'adaptation, sont en passe d'être colonisés par la généralisation de boucles de rétroaction, dont l'intelligence artificielle n'est qu'un des composants.

C'est cette bonne vieille idée du thermostat qui régule la température d'une pièce par un retour d'information et la comparaison à une norme, pour déterminer le comportement du radiateur. Imaginons donc des thermostats pour toutes nos activités sociales et économiques : du nombre de nos pas à notre activité sur les réseaux sociaux, de l'essuie-glace de la voiture (où se trouve un capteur) à la cabine des tracteurs, dans le suivi des chantiers comme des conteneurs. Partout, on mesure, on date, on géolocalise, on traite l'information, la compare à des normes pour d'éventuelles mesures correctrices : des *nudges*, des « coups de pouce » pour marcher plus, des « *likes* » pour nous encourager à publier plus, des

essuie-glaces qui tracent des cartes de pluie micro-locales, les pesticides qui sont délivrés avec précision en regard des caractères de la parcelle, des plannings de chantier optimisés selon la météo, des routes des marchandises redessinées en fonction du cours des matières.

La nouveauté est celle d'un déploiement systématique de ces types de dispositifs au cœur desquels l'apprentissage artificiel joue un rôle clé, conjugué aux systèmes d'acquisition, de traitement et de propagation des données. Comprendre l'impact de l'intelligence artificielle sur notre société, c'est d'abord en comprendre l'infrastructure et l'échelle de plus en plus vaste qui se compte en flottes de dizaines de milliards de smartphones, d'ordinateurs personnels, d'automobiles, d'objets qui collectent, traitent et propagent l'information et les décisions qui en résultent. L'enjeu de cette transformation réside dans une coordination plus efficace de la société, un management algorithmique qui associe l'échelle d'opération et l'individualisation des décisions. C'est ainsi la réalisation du projet des années 1990, celui d'une production sur mesure, et en masse, pour délivrer des biens ou services analogues en s'adaptant à toutes les situations¹.

Un marché de la coordination est en train de naître. Si les méthodes d'intelligence artificielle prennent un essor formidable, c'est d'abord parce qu'elles répondent aux besoins internes des grandes plateformes numériques et ensuite car elles constituent le socle de nouveaux services, ceux justement que la transformation numérique des entreprises demande. Pour être plus concret, pensons simplement à Uber, pour lequel l'enjeu est de fournir un service de transport dans les cinq minutes qui suivent la commande. L'intelligence n'est pas juste enfouie dans un code mystérieux, elle se nourrit de la position des quelques dizaines de véhicules proches, au moment de l'ouverture de l'application, de la connaissance de la demande locale à l'instant, et fixe un prix pour assurer l'équilibre du marché local et surtout endiguer l'engorgement. C'est un mécanisme autoadaptatif, extrêmement local et dont l'échelle de temps est de l'ordre de la minute, qui est cependant généralisé à des centaines de villes. On rêve encore d'un système analogue où les feux rouges se synchroniseraient au trafic pour le fluidifier.

L'intelligence artificielle ferait une belle utopie technoscientiste, si elle n'était accusée de troubler la société et de produire des effets négatifs,

1. Pour des articles emblématiques à cet égard, cf. Philip Kotler, « From Mass Marketing to Mass Customization », *Planning Review*, vol. 17, n° 5, 1989, p. 10-47 ; James H. Gilmore et B. Joseph Pine II, « The Four Faces of Mass Customization », *Harvard Business Review*, vol. 75, n° 1, 1997, p. 91-101.

et pervers, qui encouragent sa régulation. Presque au même rythme que ses avancées techniques, une critique sociale se développe, dévoilant que, plongée dans la société, l'IA produit des effets inattendus et parfois délétères. Au-delà de ses performances intrinsèques, elle a toutes les chances de devenir un sujet politique.

L'INFRASTRUCTURE DE LA RÉTROACTION

L'idée de feed-back irrigue le champ des théories des organisations, en posant la question de la manière dont les organisations apprennent à s'ajuster à des objectifs, mais aussi à réévaluer ces objectifs. Ce fut popularisé dans le champ du management par Chris Argyris et Donald Schön avec l'idée d'apprentissage organisationnel². Ce qui aujourd'hui est neuf est la question du tempo. Le rythme que l'on connaissait dans les organisations traditionnelles était un rythme lent, celui de la reconnaissance d'un problème, de la commande d'une étude ou d'un audit, du partage des résultats dans des cercles autorisés, et celui de décisions souvent ralenties par les considérations politiques. Un rythme qui se comptait en mois, voire en années.

41

La transformation digitale peut se résumer à l'accélération phénoménale de ce rythme, par la généralisation des boucles de rétroaction à toutes les échelles. Le « *just-in-time* » de la décennie 1990 – l'idée de livrer un bien au moment où on en a besoin afin de réduire les stocks – en est une étape essentielle, et nous sommes désormais à l'âge du temps réel. Tandis qu'il fallait des semaines pour évaluer les retombées d'une campagne publicitaire, des méthodes telles que l'« A/B Testing »³ permettent désormais d'en obtenir les résultats en quelques heures. La mesure de l'opinion, en relevant dans les contenus postés sur les réseaux sociaux les sentiments positifs ou négatifs que provoque une marque, une idée, une cause, peut se faire de minute en minute. Et ne faisons qu'évoquer, tristement, les drones qui ciblent au-dessus des champs de bataille et coordonnent les opérations.

Ces processus de feed-back se développent sur un terreau bien préparé : une infrastructure informationnelle s'est mise en place sur la planète, sédimentant des couches matérielles (câbles, satellites, serveurs...),

2. *Apprentissage organisationnel. Théorie, méthode, pratique* (1978), Bruxelles-Paris, De Boeck Université, 2001.

3. Qui consiste à comparer, de manière expérimentale et systématique, deux versions (A et B) d'un produit ou d'une page web afin de savoir laquelle est la plus performante.

logicielles (protocoles, systèmes d'exploitation, applications) et servicielles, largement interopérable et qui sert des processus complexes de traitement des données s'organisant grossièrement en quatre grandes phases.

42 1) L'*acquisition* des données passe par le déploiement de capteurs dont on espère qu'ils enregistrent, de manière valide et fiable, les variables d'intérêt. Une application de sport tentera de mesurer le nombre de pas dans une période donnée mais, la longueur du pas étant incertaine, elle pourra difficilement rendre compte de la longueur des parcours, et l'on sera alors tenté, en géolocalisant le parcours, de calculer plus exactement l'activité et de calibrer le pas. Cette activité n'est pas triviale et nécessite une idée précise à la fois de l'information que l'on souhaite produire et des finalités que l'on poursuit. Elle se généralise à tous les artefacts : ordinateurs personnels, smartphones, voitures, tableaux électriques, compteurs, mais aussi machines dans l'industrie, équipements publics, la liste est longue. Des données qui sont des nombres, des textes, des images, des signaux, en grands volumes, et large variété, disponibles immédiatement, en tout point et tout moment. C'est ce qu'on appelle le *Big Data*.

2) Le stockage et l'*archivage* des données supposent non seulement qu'on les conserve de manière organisée et structurée mais qu'on puisse les partager. La solution technique dominante semble être celle des *clouds*, même si le pair-à-pair aurait pu être une solution alternative et que la blockchain offre une solution intermédiaire. C'est désormais un marché global de 220 milliards de dollars concentré à 80 % dans les serveurs d'Amazon Web Services, de Microsoft, de Google, d'IBM Cloud, de Salesforce, d'Alibaba ou de Tencent. Les données brutes nécessitent souvent un prétraitement, pour corriger les erreurs, ajouter des métadonnées, les agréger à la bonne échelle, les anonymiser, les structurer dans des bases accessibles à des dizaines voire des centaines de millions d'utilisateurs.

3) L'exploitation des données passe par la *modélisation* : les moteurs de recherche, les moteurs de recommandation pour les sites d'e-commerce, les modèles de prévision en finance, les modèles de représentation du langage, les modèles de classification des images. C'est ce domaine qui a fait des progrès avec l'art du *machine learning* (apprentissage automatique) et dorénavant du *deep learning* (apprentissage profond). Si les plateformes de *cloud computing* (informatique en nuage) sont au premier rang sur ce segment, on y voit de nouveaux acteurs émerger, OpenAI semblant être un des plus en pointe, Hugging Face se taillant la part du

lion en distribuant la plupart des larges modèles de langage. Sans entrer dans les détails, l'ensemble de ces modèles s'appuie sur l'évaluation, à partir de vastes corpus, d'un très grand nombre de paramètres (jusqu'à des centaines de milliards), ajustables à des situations particulières – ce pourquoi on les qualifie de modèles pré-entraînés – et transférables pour effectuer d'autres tâches. Le fait notable dans ce domaine est que se construit une sorte de *lingua franca*, par les langages (Python), les lieux où l'on s'échange des idées et du code (GitHub). Une communauté de quelques millions de développeurs et de « *data scientists* », dont les interactions stimulent la créativité.

4) La *propagation* de l'information dans les systèmes de gestion est la dernière brique du processus. Elle concerne d'abord les opérations, notamment les plus répétitives : évaluer le risque d'un crédit, déterminer la réponse d'un chatbot, et dans ces cas elle contribue à l'auto-

43

matisation de la décision d'autant plus utilement que son échelle est grande. On y soulignera l'importance croissante des API, ces interfaces de programmation d'application qui assurent les échanges de données entre les plateformes : elles sont les vaisseaux sanguins qui transportent l'information et les décisions dans les tissus les plus fins de la société. La propagation peut passer aussi par la production d'éléments de connaissance pour les agents humains, c'est sa dimension épistémique. Ces éléments peuvent quant à eux prendre la forme de tableaux de bord affichés sur l'écran des managers, d'analyses d'images et de résultats de diagnostics pour les médecins, de notes de synthèse dont les IA rédigeront une part de plus en plus substantielle. C'est sur cet aspect que le développement de l'intelligence artificielle générative, celle qui se destine à produire des nouveaux documents, est susceptible de jouer un rôle très important, en automatisant la production de textes courts et contextualisés qu'assuraient les scribes de notre société : refuser une candidature à un emploi, répondre à une réclamation client, synthétiser le compte rendu d'un match, écrire un fait divers, remanier une dépêche AFP dans le style d'un journal spécifique, formuler un contrat de location standard, rédiger l'abstract d'un article scientifique, la description d'un produit sur un site d'e-commerce, et permettre ainsi des gains de productivité majeurs dans la production textuelle du monde. Le scribe automatique est désormais là, les moines copistes vont disparaître de nos économies.

La nature de la transformation digitale consiste à mettre en œuvre ces boucles de contrôle dans un modèle de management qui n'est pas fondé sur les stocks, ni sur les réseaux, mais sur un traitement de l'information

et des décisions en masse, en continu et en temps réel. Pour l'entreprise, et plus largement toutes les organisations, la signification de cette transformation réside dans sa capacité à employer ces ressources pour une production moins coûteuse et des ventes mieux valorisées. Sur un mode plus pragmatique, la transformation digitale est un impératif managérial qui vise à généraliser ces boucles dans les processus fondamentaux de ses activités : s'approvisionner, produire, vendre, maintenir.

L'idée n'est pas nouvelle. Dans les années 1990, Dell a atteint près de 25 % de parts du marché mondial avec une philosophie de gestion : le « *build-by-order* », qui ne consistait plus à prévoir un proche futur pour adapter un système de production, en réglant les erreurs par les stocks, mais à être capable de programmer la cascade des opérations en répercutant en aval les conséquences de chaque commande. Un système moins coûteux et plus adaptable. Un système flexible. La trajectoire reste la même, à une différence importante près : dans les systèmes contemporains, la disponibilité de l'information et l'amélioration des algorithmes contribueront sans doute à une plus grande autonomie des systèmes.

44

UN MARCHÉ DE LA COORDINATION

On a souvent tendance à croire que l'économie suit un fleuve tranquille jusqu'au moment où de nouvelles technologies bouleversent le cours des choses. Les technologies tomberaient du ciel comme des météores : cette thèse est celle de l'effet exogène, c'est une thèse naïve qui néglige que la compétition crée une demande d'innovation. Dans une tradition schumpétérienne, nous penchons plutôt pour la thèse endogène selon laquelle l'innovation est un bien comme les autres, et qu'on y investit si elle promet d'atteindre les buts que l'on fixe. Si elle apporte quelque chose de neuf ou de meilleur, on peut en tirer un bon prix ; pourvu que les coûts de développement ne soient pas excessifs, elle sera profitable.

L'intelligence artificielle n'est pas le surgissement d'un savoir mais l'expression d'un besoin. Pour consolider leurs positions, les géants du numérique, à chaque période de développement d'internet, ont investi pour résoudre par la technologie les problèmes qu'ils rencontraient. La fouille de données d'abord pour les besoins des moteurs de recherche (et c'est l'algorithme de Google qui a remporté la bataille), des systèmes de recommandation pour le commerce et les médias (élément clé de la stratégie d'Amazon), de l'apprentissage automatique pour les systèmes de sécurité, de sorte qu'à chaque nouveau problème répondent de nouvelles technologies. L'IA est un de ces biens.

Naturellement, les choses ne sont pas aussi schématiques, les investissements engagés ont abouti à des résultats aussi par un effet de stratification et d'accumulation : on disposait de plus grandes capacités de calcul, de plus de mémoire, de données adéquates en volume et en qualité. Les contingences scientifiques et techniques ont découvert, littéralement, des structures particulièrement efficaces. La découverte des réseaux de neurones convolutifs a ouvert la « *computer vision* », ce champ de l'IA qui s'intéresse à l'image fixe et animée et permet d'identifier les objets, de suivre une cible, d'interpréter un motif ; de même, l'idée de vectoriser les mots a trouvé, avec les « Transformers », une technique générique qui conduit depuis 2018 à la floraison des grands modèles de langage, lesquels conditionnent des progrès prodigieux en matière de traduction, de dialogue avec les chatbots, de résumé de texte, d'extraction de concept clé, de correction d'orthographe, de réponses aux questions, et désormais de génération de texte. La fin d'année 2022 a vu émerger des modèles multimodaux qui associent images, sons, textes, promettant des gains de productivité élevés dans les secteurs créatifs, du journalisme à la modélisation des protéines et jusqu'à ChatGPT, qui fait trembler les enseignants du monde entier.

45

N'oublions pas que l'information a un coût qui peut se décomposer en trois ou quatre éléments. Le coût de sa mémoire, qui a fondu avec le développement d'internet, le coût de sa recherche et de son acquisition, qui croît avec son volume (l'aiguille dans la botte de foin). Mais aussi les coûts de prédiction, que les IA tendent à réduire considérablement. C'est l'idée simple et fondamentale d'Ajay Agrawal et ses collègues : le sens économique des IA est un gain de productivité dans la production de myriades de décisions ; en réduisant les coûts de prédiction, et leur qualité, on réduit le coût des décisions en les automatisant⁴.

Juste un petit exemple. Si dans les années 2010 les commentaires des lecteurs étaient, pour les médias, une source de valeur, ils ont par la suite constitué un risque d'expression des passions mauvaises et exigent dorénavant une modération attentive mais humainement coûteuse. Un modèle tel celui du projet *Perspective* de Google, en calculant une probabilité de toxicité du contenu, peut traiter des millions de commentaires, ne laissant au soin de l'œil humain que les plus ambigus. On comprendra aisément que, par les gains de productivité obtenus, c'est toute la capacité de coordination des entreprises qui est bouleversée. En résolvant ces

4. *Prediction Machines: The Simple Economics of Artificial Intelligence*, Boston (Mass.), Harvard Business Review Press, 2018.

milliards de micro-prédictions, on automatise la décision et on améliore la coordination des organisations. (On se défait aussi d'une grande partie de la bureaucratie qui en assurait la tâche.)

Le modèle économique qui se dessine nécessite une forte interopérabilité et se réalise par la technique des API, qui distribuent les données et le calcul de manière contrôlée. Désormais, l'intelligence artificielle se vend au kilo. Comprendons bien le jeu des échanges, qui fournit la donnée, qui fournit l'intelligence, et à quel prix se fait l'échange. Un client appelle un spécialiste (une IA) qui appartient à une plateforme de données, il envoie une chaîne de caractères et souhaite un résultat : probabilité d'insatisfaction, d'une menace, etc. L'opération peut se faire des centaines, des milliers de fois chaque minute, l'IA facture en fonction du nombre de calculs. Pensons au paiement, pensons au crédit. Combien de plateformes sont mobilisées ?

Du marché paysan à l'hypermarché, des galions espagnols aux porte-conteneurs, de la forge à la sidérurgie, les marchés ont évolué en étendue, en complexité, en interdépendance, dans une organisation qui distribue le long de la chaîne de valeur des opportunités de génération de profits. Le commerçant prend le pas sur le producteur, le financier sur le commerçant, les informaticiens et les « *data scientists* » sur les financiers.

Qu'il s'agisse de détecter une mine probable, d'organiser le traitement des terres agricoles, de surveiller des troupeaux, de piloter des usines, des camions, des voitures, les boucles de rétroaction sont absolument nécessaires et leur traitement industriel se concentre. Et l'aval n'y échappe pas, si depuis les années 1990 les spécialistes de management sont convaincus de la nécessité de la personnalisation des produits, la capacité à la fabriquer se concentre dorénavant dans quelques nœuds névralgiques.

Au sein de cette chaîne de valeur, les instances de coordination deviennent elles-mêmes des chaînes de valeur indépendantes, et prélèvent leur quotité de coûts. C'est la thèse d'une nouvelle couche dans l'économie, la couche de la coordination, et d'un marché à part entière qui peut capturer 10 % à 30 % de la valeur, en la faisant croître d'autant. Voilà le facteur de leur domination : les instances de coordination apportent plus qu'elles ne prennent.

La chaîne de valeur de la coordination est de ce fait éclatée, à chacune des phases un acteur domine ou peut dominer. Le stockage des données est à présent un oligopole mondial. Ses acteurs sont les premiers candidats pour s'emparer du marché de l'IA, tirant profit des masses de données qu'ils gèrent et de leurs investissements massifs en recherche et développement (de l'ordre de 40 milliards de dollars pour Amazon, par exemple).

On conclura en s'interrogeant sur l'efficacité de ce modèle de management qui confie aux algorithmes une fraction de plus en plus importante des décisions. Il présente pour beaucoup d'acteurs des espérances importantes, à des coûts raisonnables. En matière de stratégie, il y a deux théories qui se complètent plus qu'elles ne s'opposent. L'une, celle de l'économie industrielle, dit que le succès vient quand on a trouvé la bonne position concurrentielle : celle justement qui fait échapper à la concurrence et inscrit l'activité dans un flux croissant. L'autre met l'accent sur la capacité à contrôler les ressources clés : matières premières, produits industriels, services, et désormais flux de données qui permettent un ajustement au plus près des activités. C'est donc une double bataille pour les firmes. Celle de la conquête de ces compétences qui rendent possibles des gains de productivité en s'appuyant sur l'action combinée de l'acquisition, du traitement et du déploiement de l'information, afin de favoriser l'autoadaptation des systèmes de production et de distribution. La seconde est la recherche de la meilleure position dans la chaîne de valeur : faut-il contrôler les capteurs, le calcul, la communication, l'hébergement des données ?

47

Dans le secteur automobile on en voit bien l'enjeu. La révolution de ce secteur ne passera pas par la transition électrique, mais par la capacité à utiliser les voitures elles-mêmes comme des capteurs (ce qu'elles sont déjà) et à transformer ces données (c'est le rôle des IA) en information pour des services aussi divers que l'assurance, la maintenance préventive, la gestion des flottes, l'autopartage, la prévision météo, la gestion intelligente des villes et des routes, et on en oublie.

Le monde qui se profile n'est pas forcément détestable, car c'est un monde qui promet des gains de productivité, quand le monde physique en gagne peu, et en perd même à mesure qu'on lui impute le coût de ses externalités. Mais il ne faut pas s'attendre à des miracles puisque les gains se font sur les marges, les interstices. Le modèle du management algorithmique n'est pas une révolution, c'est une propagation, celle du calcul au plus près. La croissance future passe par une économie des ressources matérielles, et la précision du calcul peut y contribuer.

LA POLITIQUE DES ALGORITHMES

Pour le lecteur critique, la société du nombre peut effrayer et révolter. On comprend que, pour le romantique, le sujet insécable et autonome des Lumières se confronte à un casse-tête : celui d'être dépendant des

chiffres qui produisent ses cibles. La critique de la commensuration d'Alain Supiot va dans ce sens⁵.

La critique sociale de cette algorithmisation du monde s'appuie sur l'observation ou le soupçon de certains effets sociaux pervers. Pour les juristes, les systèmes issus de ce processus possèdent un caractère de boîte noire pas forcément compatible avec l'exigence de transparence⁶. Très rapidement les effets discriminatoires qui renforcent les stéréotypes sociaux ont été mis en évidence, l'hypothèse de soupçon de bulles de filtres qui enferment les usagers des réseaux sociaux a été partiellement validée, l'amplification des désinformations ajoute à la détermination de bandes organisées, et la multiplication d'algorithmes voués à la mise en avant de contenus tendance déforme le miroir du monde.

48 Cette critique sociale peut prendre une dimension moraliste et s'associer à un néo-luddisme⁷. Elle alimente l'effort de régulation des autorités et renforce des principes fondamentaux de respect de la vie privée – pour les algorithmes, la nécessité de rendre compte –, pouvant aboutir à l'interdiction de certaines mesures, et de certains calculs : la question de la reconnaissance faciale se pose, l'obligation de déclarer qu'un texte a été rédigé par une machine appartient à un futur très proche.

La critique sociale a l'intérêt de mettre en évidence un phénomène essentiel : quand les boucles de contrôle quittent le monde physique (le radiateur) pour rejoindre le monde social (les foules), les résultats ne sont plus déterminés ; quand le numérique se confronte au social, les résultats ne sont jamais ceux qu'on attend, la foule en détourne l'usage et la mesure, et l'instrumentation doit être redressée.

Le cas du système de réputation d'eBay est remarquable. Ce système était jusqu'en 2009 symétrique, le vendeur et l'acheteur se notant mutuellement, il favorisait la rétorsion (quand l'acheteur punissait le vendeur par une mauvaise note, le vendeur faisait de même), entraînant une dégradation des notes moyennes, et par conséquent une moindre stimulation des ventes. Il fut nécessaire de changer de thermomètre. La note

5. *La Gouvernance par les nombres*, Paris, Fayard, 2015.

6. Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (Mass.), Harvard University Press, 2015.

7. Parmi les conséquences sociales du déploiement des IA, il faut *a minima* mentionner les répercussions sur le marché du travail selon l'hypothèse de polarisation formulée par David Autor et David Dorn : si les IA amélioreront la position et la productivité des classes créatives, l'automatisation qu'elles impliquent touchera les classes intermédiaires en se substituant à leurs membres, ce qui augmentera la compétition parmi les moins qualifiés (« The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market », *American Economic Review*, vol. 103, n° 5, p. 1553-1597).

réciroque est devenue la fréquence des transactions satisfaisantes, qui tutoie en général les 98 % et apporte ainsi un élément de confiance à l'ensemble de la population. Dans ce cas, l'algorithme est élémentaire.

Exemple plus récent : le projet *Galactica*, lancé par Meta en novembre 2022 et qui, après trois jours de fonctionnement, a fermé son accès au public. Le risque était trop grand de voir la génération de texte se mettre au service des propagandes, de la désinformation, des publications délirantes. L'outil puissant ne possédait pas de mode d'emploi social. Qui peut l'utiliser, pourquoi ? Avec quel risque ?

C'est que nous aimons appeler « politique des algorithmes » ce pouvoir qui affecte la société quand le calcul échappe à ses concepteurs et est « recalculé » par ses utilisateurs. Politique des algorithmes car ils exercent un pouvoir en redistribuant le pouvoir. S'ils n'ont pas de volonté, et certainement pas celle de dominer, par le jeu subtil de la distribution des pouvoirs ils acquièrent un pouvoir qui se révèle dans l'interaction des boucles de contrôle algorithmiques et des populations qui peuplent nos sociétés, raison suffisante pour en faire un acteur au centre des sciences sociales.

49

Les algorithmes affectent aussi la nature de la décision, et posent d'eux-mêmes un problème politique. Dans des processus décisionnels lents, la phase politique ou délibérative est essentielle, les éléments du calcul sont interprétés, discutés, contextualisés, et la question clé est celle de la gouvernance : qui participe à la décision avec quels droits ? Quand la décision est algorithmique, le contexte politique s'efface, la décision est une anticipation, et ce sont ses erreurs qui sont délibérées. On passe donc d'un jugement *a posteriori* à un jugement *a priori*, qui nécessite son complément, celui de la correction d'erreurs (les services de réclamation client ont un grand avenir). La présomption d'innocence cède la place à la présomption de culpabilité. Dans le modèle lent la décision est prise après avoir été mise au gril de sa justesse, dans le modèle rapide elle est prise au risque de l'erreur, celle des faux positifs et des faux négatifs.

*

Pour conclure, attachons-nous à cette inquiétude sourde : la démocratie peut-elle être cybernétique ? Oui, pour autant que ses machines ne soient pas des boîtes noires, que les citoyens en aient une connaissance suffisante, que ces systèmes soient le plus ouverts possible et puissent être contestés. Plus encore, leurs utilisateurs doivent pouvoir en contrôler les effets, voire les débrancher. Mais par-dessus tout, bien que ce soit

utopique, de la même manière que la loi se forge dans le débat public, les sujets des algorithmes devraient pouvoir intervenir dans la construction des systèmes, le choix de leurs architectures et de leurs paramètres (notamment leur taux d'erreur). C'est finalement la question posée par les foucades d'Elon Musk et de ses détracteurs, à laquelle aucune arène constituée ne pourra répondre, sauf celle des rapports de force et des forces du marché.

50 Les IA ne sont pas de pures machines et ne peuvent pas seulement être évaluées dans les termes techniques de leur efficacité prédictive, souvent comparée à la performance humaine, qui est déjà dépassée dans de nombreux domaines. Les IA, par leur incorporation dans les dispositifs algorithmiques, sont des acteurs sociaux. Leur existence réside dans leur acceptation sociale, leur légitimité. Le paradoxe de la société du calcul est que la société tend à déjouer les résultats du calcul. La vertu du calcul n'est pas dans sa fiabilité mais dans sa capacité à installer un ordre accepté.

R É S U M É

Pour comprendre la manière dont les intelligences artificielles affectent la société, il est nécessaire, d'abord, de les situer dans un ensemble plus vaste de techniques et de systèmes qui déploient sur la société une infrastructure encourageant le développement de boucles de contrôle des activités sociales et économiques. Ensuite, il faut se rappeler qu'elles sont le fruit d'une économie à la recherche constante de gains de productivité ; elles constituent de ce point de vue le produit endogène de la plateforme du monde. Finalement, parce qu'elles sont plongées dans la société, leurs effets inattendus, parfois délétères, en font des sujets politiques. L'enjeu essentiel est celui de leur légitimité.